English-Language Arts - Writing Tasks

SCORING PROCESS

Questions and Answers

Who scores the student responses and how are they selected?

The people who scored the student responses are college graduates who possess at least a BA degree. Whenever possible, educators are hired as readers (scorers).

Are California teachers involved in the scoring process?

Yes, California teachers are involved in several aspects of the scoring process. They participate in the field testing of new CAHSEE writing prompts, help select sample student essays to be used for training, serve as readers (scorers), and serve as scoring leaders.

How are the readers (scorers) trained to score student responses?

Each prospective reader is required to take an extensive on-line tutorial.

The tutorial consists of the following:

- General information about the Online Scoring Network (OSNTM) system
- Background information about the California ELA standards
- Extensive information on the CAHSEE item types
- Explanations of CAHSEE scoring guides and scoring principles
- Detailed topic notes on selected topics
- Sets of annotated training papers for these selected topics. The training materials include Benchmarks and Rangefinders. Benchmarks are solid

samples of student writing for each score point (1-4). Rangefinders are samples of student writing which demonstrate the "high" and "low" end of each score point.

The tutorial also includes a practice "certification test."

How does a potential scorer QUALIFY as a reader?

After completing the CAHSEE OSN Tutorial, readers must pass an official certification test (a total of 40 papers, 20 from each item type) before being eligible for consideration as readers for the CAHSEE program. To pass, readers need to score at least 70% exact and have no discrepancies (more than one point apart from the predetermined score).

Before each scoring session, each reader must pass a Calibration test of 10 papers (from the item type he or she is scheduled to read during that session) with seven exact scores and with no discrepancies (more than one point away from the predetermined score). A reader has up to two opportunities on three discrete calibration tests, with retraining by the scoring leader between each set. If a reader does not pass one of these sets, he or she is not allowed to score.

English-Language Arts - Writing Tasks

SCORING PROCESS

How are Scoring Leaders selected?

The most experienced readers with the highest accuracy rate on the certification test are selected to be scoring leaders. Scoring leaders monitor and mentor readers during operational scoring. There are approximately six readers assigned to one scoring leader; this allows scoring leaders to work closely with each reader. The OSN system also allows scoring leaders to read essays simultaneously with readers so that they easily can work with readers on individual papers.

The majority of scoring leaders participate in a one-day training session. This training covers the California standards, the scoring guides, the operational topics, and a wide variety of sample papers. The session also includes detailed information on how to work with readers and how to use the OSN system. During operational scoring, additional newly-selected scoring leaders are individually trained by an experienced scoring leader.

How are the training papers chosen?

Working with experienced readers, content specialists review the student responses from the field tests to look for representations of each of the score points and the variety of approaches used to respond to the prompt. They choose samples that illustrate the criteria in the scoring guides and the possible ways in which students may have responded to the topic. The responses are then put into sets used for training and qualifying readers.

Are the responses scored on computers?

Yes. All of the student writing responses are scanned into the state-of-the-art OSNTM system. (The OSN system is currently used to score such high-stakes writing assessments as the GMAT-Analytical Writing Assessment, the GRE Writing Measure, and such programs as TOEFL and CLEP.) The readers view each response via the Internet on individual home computer screens on which they assign the scores. The screen does NOT display any background information about the student; the reader sees only the actual essay. Scoring leaders monitor readers from individual home computers, also via the Internet.

What sort of scoring guide do readers use to evaluate student responses?

The CAHSEE writing task can either be a "Response to Literature/Text" task or a "Response to Writing Prompt" task.
Readers use a different scoring guide for each type of writing task.

The scoring guide used for a "Response to Literature/Text" task includes the criteria in the ninth and tenth grade California Content Standards for "Writing Applications" for standard 2.2 – writing a response to literature. This guide reflects the need for textual support in order to earn the higher scores. The texts may be literary, including dramas, stories, poems, and literary nonfiction (e.g., essays, autobiographies, biographies, or memoirs written in a literary style). The texts also may be informational, persuasive, or functional texts (e.g., consumer materials or workplace documents).

English-Language Arts - Writing Tasks

SCORING PROCESS

All texts on the CAHSEE are carefully screened so that they are free from potentially biased, sensitive, or controversial content.

The scoring guide used for a "Response to Writing Prompt" task includes the criteria within the ninth and tenth grade for the following "Writing Applications" content standards:

- ➤ 2.1—writing biographical compositions
- ➤ 2.3—writing expository compositions
- ➤ 2.4—writing persuasive compositions

This guide reflects the skills needed to write a composition that is <u>not</u> related to a previously read text.

The scoring guides are included as part of the ELA Teacher Guide, which is available through the CAHSEE website at http://www.cde.ca.gov/statetests/cahsee/.

How many times is each student response scored?

Scores are assigned independently by each reader. Each response is scored twice, once by one reader and once by another reader, and the average of the two scores constitutes the reported score.

If the two scores are non-adjacent (more than one score point apart), the response is assigned a third score by a scoring leader. In these cases, the reported score is determined by averaging the two closest scores. If the third score falls in between the first two scores, then all three scores are averaged.

If a paper is difficult to score and needs the expertise of the scoring leader to determine the reported score, the reader defers the paper to the scoring leader to be scored.

How many points can a student response earn, and how are the scores reported?

The student response can earn up to four points. Since each essay is rated by two readers, the scores are averaged. The averaged score for each essay is reported to the student and the schools. For example, if a student's response receives a 4 and a 3, the scores are averaged and the reported score is 3.5.

ELA scores are reported on a scale ranging from 250 to 450. The students' score on the writing task is "weighted" to account for 20 percent of their ELA scale score. Students' multiple-choice scores for reading and writing are summed and then weighted to account for the remaining 80 percent of their scale score.

How is the accuracy of scoring maintained throughout the scoring process?

The quality of all scoring is monitored on a regular basis.

First, the consistency between readers is calculated based on whether the first and second readers assigned identical, adjacent, or non-adjacent scores. Scoring leaders and test developers constantly monitor agreement percentages. If a reader's rate of agreement begins to get too low, the reader is retrained by the scoring leader and closely monitored thereafter. If the reader's performance does not improve, that reader is released.

Second, ten percent of the papers that are read by each reader have been previously scored by content specialists and scoring

English-Language Arts - Writing Tasks

SCORING PROCESS

leaders. The consistency of the readers' ratings with those of the predetermined scores is checked several times a day to validate that each reader (scorer) applies the criteria in the scoring guides accurately to each student response.

Third, the scoring leaders back-read and formally record the scores for a percentage of the scored papers to ensure reliability throughout the scoring process. "Back-reading" is the process of reading papers that have already been rated by a reader, to verify the consistency of the readers' ratings.

Fourth, to help prevent readers from drifting, readers are required to pass a calibration test before each reading session. If a reader is deficient on any of the quality indices, he or she is immediately retrained or released from the scoring process.

What are the guidelines for sensitive topics on the CAHSEE?

To keep the test free from potentially biased, sensitive, or controversial content, the following topics are avoided on the examination:

- Violence (including guns, other weapons, and graphic animal violence)
- Dying, death, disease
- War
- Natural disasters with loss of life
- Drugs (including prescription drugs), alcohol, tobacco, smoking

- Junk food
- Abuse, poverty, running away
- Divorce
- Socio-economic advantages (e.g., video games, swimming pools, computers in the home, expensive vacations)
- Sex
- Religion
- Complex discussions of sports
- Slavery
- Evolution, prehistoric times, age of solar system, dinosaurs
- Rap music, rock concerts
- Extrasensory perception, witchcraft
- Halloween, religious holidays
- Anything disrespectful, demeaning, moralistic, chauvinistic
- Children coping with adult situations or decisions; young people challenging or questioning authority
- Mention of individuals who may be associated with drug use or with advertising of substances such as cigarettes or alcohol
- Losing a job, being fired
- Rats, roaches, lice, spiders
- Dieting, other concerns with self-image
- Any topic that is likely to upset students and affect their performance on the rest of the test

It is important to note that these guidelines are applied in the context of the purpose of the test as well as the overall passage or item. For example, some subjects (e.g., the animals listed above) may be mentioned in a text, although an entire passage would not focus on these subjects.